

Cross-talk and interference can enhance information capacity of a signaling pathway

Sahand Hormoz*

Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 91306, USA

(Dated: February 23, 2012)

A recurring motif in gene regulatory networks is transcription factors (TFs) that regulate each other, and then bind to overlapping sites on DNA, where they interact and synergistically control transcription of a target gene. Here, we suggest that this motif maximizes information flow in a noisy network. Gene expression is an inherently noisy process due to thermal fluctuations and the small number of molecules involved. A consequence of multiple TFs interacting at overlapping binding sites is that their binding noise becomes correlated. Using concepts from information theory we show that a signaling pathway transmits more information if 1) the noise of one input is correlated with that of the other, and 2) the input signals are not chosen independently. In the case of TFs, the latter criterion hints at up-stream cross-regulation. We explicitly demonstrate these ideas for the toy model of two TFs competing for the same binding site. We suggest that this mechanism potentially explains the motif of a coherent feed-forward loop terminating in overlapping binding sites commonly found in developmental networks, and discuss three specific examples. The systematic method proposed herein can be used to shed light on TF cross-regulation networks either from direct measurements of binding noise, or bioinformatic analysis of overlapping binding sites.

I. INTRODUCTION

Accurate transmission of information is of paramount importance in biology. For example in the process of embryonic development, crude morphogen gradients need to be translated into precise expression levels in every cell and sharp boundaries between adjacent ones [1, 2]. The embryo accomplishes this using a complex network of signaling molecules that not only regulate the expression level of the desired output gene, but also each other. One simple strategy for increasing accuracy is the use of multiple input signals. Indeed, frequently, the expression level of a single gene is controlled by multiple transcription factors (take for example *bicoid* and *hunchback*, or *dorsal* and *twist* in the *Drosophila* embryo [1, 3, 4]). These transcription factors, however, often have overlapping binding sites that result in interactions at binding and synergetic control of transcription [3, 4].

Here, we suggest that interaction at the level of binding (interference) is related to the upstream network of transcription factors regulating each other (cross-talk). Our main assumption is that the regulatory network is designed to optimize information transfer from the input (TF concentrations) to the output (gene expression level). This is a reasonable assumption in the case of development, where accurate positional information needs to be extracted from noisy morphogen concentrations [2].

First, we define the concept of a cis-regulatory network as a noisy communication channel, where the input encodes information by taking on a range of values, i.e. a morphogen gradient that carries positional information. Decoding this information is subject to biological noise; for example, at the molecular level, the stochastic binding of morphogens to receptors makes an exact read-out of their concentration impossible.

We show that in general two input signals with correlated read-out noise can transmit more information if they are not independent information carriers but chosen from an ‘entangled’ joint-distribution, i.e. the concentration of one morphogen in a given cell is related to that of the other. Physically, this implies that the two inputs regulate each other upstream through cross-talk. We demonstrate this for the simplest possible case by writing a toy model of two TFs competing for the same binding site. The competition at the binding site results in correlated binding/unbinding fluctuations. Solving for the optimal joint-distribution of the input TF concentrations indicates that upstream, one is positively regulated by the other. Despite the increase in noise for each individual input from the competition, two interacting TFs can transmit more information than two non-interacting TFs because of 1) correlated noise in input read-out, 2) an entangled optimal input distribution.

We suggest that this mechanism is consistent with the recurring strategy of the coherent feed-forward motif, where one TF positively regulates another, and both bind to partially overlapping sites that induce interactions. In particular, we discuss the joint regulation of gene *Race* in the *Drosophila* embryo by intracellular protein Smads and its target *zen*; regulation of *even-skipped* stripe 2 by *bicoid* and *hunchback*; and that of *snail* by *dorsal* and *twist*. Generalization to other forms of cross-talk, such as cross-phosphorylation, and other forms of interference, such as use of scaffold proteins, is also discussed.

Another purpose of this work is to communicate the broader importance of correlated noise in biological systems. The conventional approach of treating intracellular noise as uncorrelated fluctuations might mask the functional importance of intricate network structures used in replication, gene regulation, transcription, and translation.

* hormoz@kitp.ucsb.edu

II. GENE REGULATION AS A COMMUNICATION CHANNEL

Regulatory networks in a cell are information processing modules that take in an input, such as concentration of a nutrient, and generate an output in the form of a gene expression level. Information in the input is typically encoded as the steady-state concentration of a transcription factor c , which binds to the promoter site of the desired response gene and enhances or inhibits its transcription. At a molecular level, the process of binding is inherently noisy, subject to thermal agitations and low-copy number fluctuations [5–7]. The noise is captured through a probabilistic relationship between the binding site read-out n , and TF concentration c , $p(n|c)$. Detailed form of $p(n|c)$ depends on physical parameters such as binding and unbinding rates. We can think of this process as communication across a noisy channel where read-out n from the binding site does not correspond to exactly the input c . To alleviate impact of the noise, various strategies can be adopted, such as limiting the input to sufficiently spaced discrete concentration levels c_i that result in non-overlapping read-out distributions n_i .

Shannon’s channel coding theorem [20, 21] tells us the maximum rate at which information can be communicated across a noisy channel, or the channel capacity. Throughout this work, we will assume that gene regulatory networks are selected to optimize the rate of information transmission. This is a strong but reasonable assumption, for example, the cell will clearly benefit from a more accurate knowledge of the amount of nutrient in its environment. However, the cost of an optimal networks can exceed the benefit of more accurate information. In this work, we do not account for the cost of a network, the only metric for comparison is the channel capacity.

With knowledge of the nature of the noise in a channel $p(n|c)$, it is possible to compute the probability distribution of the input signal $P_{TF}^*(c)$ that maximizes rate of information transmission. Essentially, this distribution tells the sender how often a particular TF concentration should be used for optimal transmission of information encoded in concentration (take for example spacing of discretized inputs suggested above). However, it does not tell the sender anything about the encoding and decoding schemes. This abstraction is useful, allowing us to compute the optimal input without having derived the optimal coding. However, the optimal coding might require input blocks of infinite size and complex codebooks, with little biological relevance.

Nevertheless, there are experimental observations consistent with the idea of regulatory systems maximizing information transmission rates. Tkacik et al. [8] have shown that experimental measurements of Hunchback concentration in early *Drosophila* embryo cells [12] has a distribution that closely matches the optimal frequency for the measured levels of noise in the system; with the system achieving 90% of its maximum transmission rate. Even for systems that are not optimized, channel capac-

ity is a useful metric for comparison. A particular architecture of transcription factors and binding sites can potentially achieve a higher transmission rate than others.

We consider the case of two transcription factors encoding information in their concentrations $c_{1,2}$ for regulation of a desired gene. First, we show that correlations in the read-out noise of the two transcription factors improves channel capacity. This is not surprising. Next, we write down a simple physical model of generating correlated read-out noise: competition for the same binding site. Competing transcriptional modules have been extensively observed in both prokaryotes [9] and Eukaryotes [10]. Competing for a binding site clearly increases noise in read-out of each TF individually. What is not clear a-priori is whether the gain from correlations can compensate for loss in accuracy of the individual read-out. We show that surprisingly competing transcription factors can transmit information at a higher rate.

Essentially, this follows from fully exploring the optimal joint-distribution of the frequency of inputs $P_{TF}^*(c_1, c_2)$, which is no longer separable $P_{TF}^*(c_1, c_2) \neq P_1^*(c_1)P_2^*(c_2)$ in the case of competing TFs. More importantly, $P_{TF}^*(c_1, c_2)$ reveals correlations between the input TFs, shedding light on their upstream interactions. With correlated noise at the read-out level, the channel is optimized when the two transcription factors regulate each other, and are no longer independent information carriers. Below, we find that one TF ought to positively regulate the other. Comparison to biological system reveals a predominance of the ‘feed-forward’ motif for transcription factors that compete for the same binding site, not inconsistent with our results.

III. NOISE CORRELATIONS ENHANCE CAPACITY

First, we quantify how correlations in read-out noise enhance rate of information transmission, following closely the approach of [11]. Consider two transcription factors with concentrations c_1 and c_2 that regulate expression level of a gene (denoted as g) through receptor binding readouts $n_{1,2}$. To carry information $c_{1,2}$ must take on a range of values. These values can vary for example as a function of space, as in the case of morphogens along an embryo. The frequency of observing a particular concentration occurrence c_1 and c_2 is given by $P_{TF}(c_1, c_2)$. The information content is maximized when this distribution is uniform, or all concentrations equally likely. Of course, our aim is not to maximize information content in $c_{1,2}$, but rather the information conveyed to the expression level g .

For now, we neglect the details of the read-out mechanism (c to n) and focus on regulation control from $c_{1,2}$ to g . The noise in the expression levels results in a distribution of g for fixed TF concentrations, $P(g|c_1, c_2)$. Equivalently, we can fix the expression level g and consider the

corresponding distribution of TFs, $P(c_1, c_2|g)$, assuming that there is unique set of inputs for every value of g . The two distributions are related by Bayes' rule. The amount of information communicated from $c_{1,2}$ to g is given by the mutual information between the distributions of $c_{1,2}$ and g [20],

$$I(g; c_1, c_2) = - \int dc_1 dc_2 P_{TF}(c_1, c_2) \log P_{TF}(c_1, c_2) \dots \\ + \int dg P_{exp}(g) \times \int dc_1 dc_2 P(c_1, c_2|g) \log P(c_1, c_2|g) \quad (1)$$

where the distribution of expression level g is given by $P_{exp}(g) = \int dc_1 dc_2 P(g|c_1, c_2) P_{TF}(c_1, c_2)$.

We assume that the noise in $c_{1,2}$ for a fixed expression level g is small and distributed as a Gaussian around the mean value $\bar{\mathbf{c}}(g)$,

$$P(c_1, c_2|g) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp \left[-\frac{1}{2} (\mathbf{c} - \bar{\mathbf{c}}(g))^T \Sigma^{-1} (\mathbf{c} - \bar{\mathbf{c}}(g)) \right], \quad (2)$$

where $\mathbf{c} = (c_1, c_2)$, and Σ is the covariance matrix over the conditional probability for fixed g , or the noise covariance matrix,

$$\Sigma_{ij}(g) = \langle (c_i - \bar{c}_i(g))(c_j - \bar{c}_j(g)) \rangle. \quad (3)$$

Essentially, the small-noise approximation says that it is meaningful to think of a mean input-output response, which is what is commonly measured in experiments. We expand around the mean response to the next order. The approximation, although strong, has been verified for a variety of regulatory systems (see for example Bicoid-Hunchback in [12], or for other examples [13–15]), and enables us to analytically calculate the optimal distribution. The mutual information under this approximation is given by,

$$I(g; c_1, c_2) = - \int dc_1 dc_2 P_{TF}(c_1, c_2) \log P_{TF}(c_1, c_2) \dots \\ + \frac{1}{2} \int dc_1 dc_2 P_{TF}(c_1, c_2) \log \left(\frac{|\Sigma^{-1}(\bar{\mathbf{g}}(\mathbf{c}))|}{4\pi^2 e^2} \right) \quad (4)$$

where Σ^{-1} is evaluated at the mean value of expression level \bar{g} corresponding to a given \mathbf{c} .

To find the channel capacity, Eq.(1) is optimized for the input distribution $P_{TF}(c_1, c_2)$. We constrain the input concentration to lie in the range $c_{1,2} \in [c_{min}, c_{max} = 1]$, where the maximum concentration is normalized to one. The minimum concentration is set by the molecular nature of the input: a minimum of one input molecule per cell is required. With the probability distribution's normalization constraint introduced using a Lagrange multiplier, the optimal distribution must satisfy,

$$\frac{\delta}{\delta P_{TF}(c_1, c_2)} \left[I(g; c_1, c_2) - \lambda \int dc_1 dc_2 P_{TF}(c_1, c_2) \right] = 0. \quad (5)$$

The optimal input distribution in the small-noise approximation (Eq.(4)) is given by,

$$P_{TF}^*(c_1, c_2) = \frac{1}{2\pi e Z} \frac{1}{\sqrt{|\Sigma|}}, \quad (6)$$

where Z is the normalization constant.

The maximum mutual information, or channel capacity for transmitting information from TF concentrations to expression level equals,

$$I^* = \log_2 Z = \log_2 \left[\frac{1}{2\pi e} \int \int_{c_{min}}^1 dc_1 dc_2 \frac{1}{\sqrt{|\Sigma|}} \right]. \quad (7)$$

We can repeat the same computation for one TF while neglecting the other, effectively ignoring the covariance of the noise (off-diagonal components of Σ). With no covariance, the noise distribution is separable, $P(c_1, c_2|g) = P(c_1|g)P(c_2|g)$. The optimal input concentration for TF₁ will be $P_1^*(c_1) \sim \frac{1}{\sqrt{\Sigma_{11}}}$, and its channel capacity, $I_1^* \sim \log \int dc_1 \frac{1}{\sqrt{\Sigma_{11}}}$; with a similar expression for the other TF.

For the simple case where Σ is independent of \mathbf{c} , channel capacity of the two TFs can be decomposed into its individual and joint contributions,

$$I^* = I_1^* + I_2^* - \frac{1}{2} \log(1 - \rho^2), \quad (8)$$

where $I_{1,2}^*$ is the channel capacity of the transcription factors individually, and $\rho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}$ is the noise correlation coefficient for TF concentrations. Accounting for noise correlation enhances the rate of information transmission. In fact, in the limit of perfect correlation, $\rho \rightarrow \pm 1$, the capacity is infinite. This is expected, since under the small-noise approximation and perfectly correlated noise, some combination of inputs is always noise free. Noise-free continuous variables can transmit infinite information.

In general, the optimal input distribution with noise-correlation is not separable to individual components, namely,

$$P_{TF}^*(c_1, c_2) \neq P_1^*(c_1)P_2^*(c_2), \quad (9)$$

where $P_{1,2}^*$ is the marginal distribution for $c_{1,2}$. In a sense, $P_{TF}^*(c_1, c_2)$ is an entangled distribution, where the concentration of one TF determines the probability of observing a certain concentration of the other. Biologically, this hints at upstream interactions between the transcription factors; the form of which should be predictable from the nature of the noise correlations.

The above abstract results are not surprising. The more important question is whether noise can be correlated, i.e. $P(c_1, c_2|g) \neq P(c_1|g)P(c_2|g)$, for the physical process of binding and unbinding of multiple TFs to a promoter region. We will demonstrate this below using competing transcription factor modules.

IV. COMPETING TRANSCRIPTION FACTORS

Transcription factors regulate gene expression levels by binding to cis-regulatory regions on the DNA. The design of these regions is highly complex in both prokaryotes and eukaryotes, with overlapping TF binding sites occurring frequently [9, 10, 30].

We write a simplified model of the two extremes of overlapping binding sites (Fig.1). Fig.1A depicts two TFs interacting independently with their corresponding binding site, whereas, Fig.1B depicts the two TFs competing for the same site. In a gross simplification, we assume that the dominant source of noise is from fluctuations in binding/unbinding of TF to the promoter. Details of RNAP assembly and transcription are coarse grained to a simple TF binding picture. The discrete nature of output and its stochastic degradation, along with fluctuations in input concentrations are discarded. Nonetheless, we will show that this simple model captures the essential role of noise correlations in a regulatory network.

Following the approach of [17], let $n_{1,2}$ be the fractional occupation of the binding site by TF_{1,2}. $n_1 + n_2 < 1$ is the fractional occupation of the site by either TF. A binding event can occur only if the site is unoccupied, $1 - n_1 - n_2$ of the time.

$$\begin{aligned} \frac{dn_1(t)}{dt} &= k_1 c_1 (1 - n_1 - n_2) - l_1 n_1 \\ \frac{dn_2(t)}{dt} &= k_2 c_2 (1 - n_1 - n_2) - l_2 n_2 \end{aligned} \quad (10)$$

The binding rate (on-rate) is proportional to the concentration of TF present, and the off-rates given by constants $l_{1,2}$. At thermal equilibrium these two rates, are related through the principle of detailed balance, $\frac{k_1 c_1}{l_1} = \exp\left(\frac{F_1}{k_B T}\right)$, where F_1 is the free energy gain in binding for TF₁, with a similar expression for the other TF. Throughout, this section we rescale time so that $k = 1$. As per previous section, TF concentrations are measured in units of $c_{max} = 1$.

Eq.(10) is a dynamical picture of the fractional occupation of the binding site by each TF. At steady state the mean fractional occupation is denoted by $\bar{n}_{1,2}$. We incorporate thermal fluctuations around this mean by introducing small fluctuations in the rate constants $\delta k_{1,2}$ and $\delta l_{1,2}$, and linearizing fluctuations in $\delta n_{1,2}$ around the mean. The TF concentrations do not fluctuate; they are the fixed inputs of the system. Fluctuations in $n_{1,2}$ effectively introduce noise in the estimates of the concentrations, $p(\mathbf{n}|\mathbf{c})$.

From detailed balance, we know that δk and δl fluctuations around steady state can be replaced by the fluctuations in free energy δF alone. With this substitution and taking the Fourier transform, the linearized equations take the form,

$$\begin{pmatrix} \delta \tilde{F}_1 \\ \delta \tilde{F}_2 \end{pmatrix} = \frac{k_B T}{1 - \bar{n}_1 - \bar{n}_2} \begin{pmatrix} 1 + \frac{-i\omega + l_1}{k_1 c_1} & 1 \\ 1 & 1 + \frac{-i\omega + l_2}{k_2 c_2} \end{pmatrix} \begin{pmatrix} \delta \tilde{n}_1 \\ \delta \tilde{n}_2 \end{pmatrix}$$

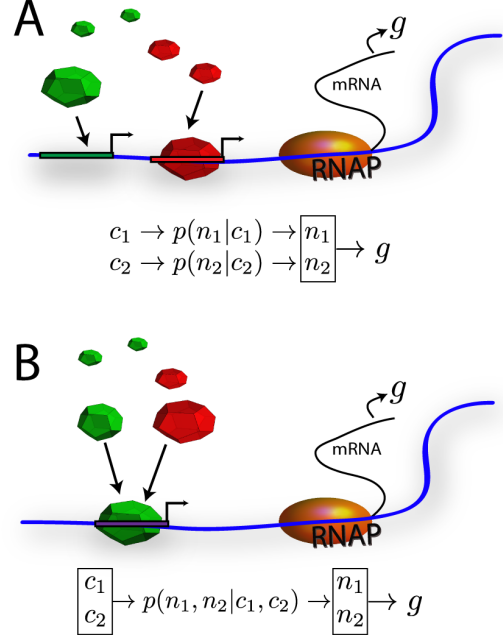


FIG. 1. Independent vs. competing transcription factors. A) Non-overlapping binding sites. The read-outs $n_{1,2}$ are not correlated, so is the noise in estimating $c_{1,2}$. The expression level g is dependent on both inputs. B) Over-lapping binding sites. The read-outs are dependent, resulting in correlated noise in estimating c_1 and c_2 .

where tilde denotes the Fourier-transform, i.e. $\delta \tilde{n}_1(\omega) = \int_0^\infty \delta n(t) e^{i\omega t} dt$. In vectorial form, the relation becomes, $\delta \tilde{\mathbf{F}} = \Lambda \delta \tilde{\mathbf{n}}$.

Eq.(11) relates incremental fluctuations in read-out $\delta \mathbf{n}$ with fluctuations in free energy $\delta \mathbf{F}$. This is a linear response relation, with the free energy playing the role of the driving force (for details, see [17]). Using fluctuation dissipation theorem [18], we calculate the power-spectrum of noise in \mathbf{n} .

$$S_n(\omega) = \frac{2k_B T}{\omega} \Im(\Lambda^{-1}), \quad (12)$$

with \Im denoting the imaginary part of the inverse of Λ matrix computed above. From S , we can compute the covariance matrix,

$$\langle \delta \mathbf{n}^T \delta \mathbf{n} \rangle = \int_{-1/\tau}^{1/\tau} \frac{d\omega}{2\pi} S_n(\omega). \quad (13)$$

τ denotes the integration time of the site. The read-out noise decreases with an increase in integration time. In our calculations, we assume very short integration time, namely $1/\tau \rightarrow \infty$. The off-diagonal element of $\langle \delta \mathbf{n}^T \delta \mathbf{n} \rangle$ is the covariance between the two read-outs. With proper normalization, we can compute the correlation coefficient (Fig. 2A). The correlation-coefficient is negative, since a more than expected occupation of the site by one TF

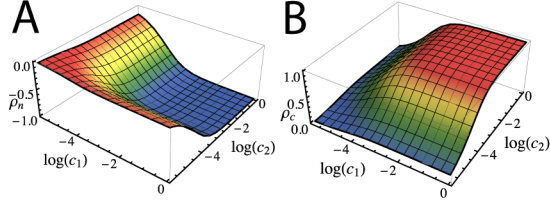


FIG. 2. Correlation coefficients. (A) Correlation coefficient of readouts n_1 and n_2 , for $l = 10^{-4}$ and $c_{min} = 10^{-3}$ as a function of log input TF concentration. (B) Correlation coefficient of the TF concentration estimates for the same parameter values.

will clearly result in less than expected occupation by the other.

Finally, we need to relate the noise in the read-out $\delta \mathbf{n}$ to the noise in the estimated TF concentrations. To do so, we account for the sensitivity of the read-out to the TF concentrations. For example, a very large c_1 results in $n_1 = 1$ with little noise. This readout, however, is not very sensitive to changes in c_1 , and not useful in detecting concentration changes. Define the matrix, $\Omega_{ij} = \frac{\partial c_i}{\partial n_j}$. The covariance matrix for the noise in TF concentrations is given by,

$$\Sigma = \langle \delta \mathbf{c}^T \delta \mathbf{c} \rangle = \Omega \langle \delta \mathbf{n}^T \delta \mathbf{n} \rangle \Omega^T. \quad (14)$$

In equating the covariance matrix in TF concentration to Σ (covariance matrix for a fixed g) of the previous section, we have introduced the extra assumption that the dominant noise in the channel going from \mathbf{c} to g is from the read-out noise and not the expression-level. Noise in g is assumed negligible and need not be propagated backwards and included in Σ . Since noise in g is most commonly shot-noise, this assumption is reasonable when expression-levels are high. This also means that our results will not depend on the functional form of g on \mathbf{c} (for the case when they do for one input see [11, 16]). Fig.2B shows the correlation coefficient for the noise in TF concentrations. The correlation coefficient is now positive.

Assuming the read-out noise computed above is the dominant source of noise, we can compute the optimal joint-distribution of input concentration $P_{TF}^*(c_1, c_2)$ by plugging the covariance matrix in Eq.(6) (Fig. 3A). Moreover, Eq.(7) tells us the channel capacity, or the maximum information transmission rate. Fig. 3B plots channel capacity of two interacting TFs and two independent ones as a function of logarithm of off-rate $\log_{10}(l)$. The interacting TFs have a higher channel capacity in the biologically relevant regime where $l \sim 10^{-4}$ and $c_{min} \sim 10^{-3}$ (see below).

The channel capacity is higher for interacting TFs, despite an increase in the noise of individual readouts, because the optimal joint-distribution of input concentrations (Fig. 3A) is ‘entangled’ and no longer sep-

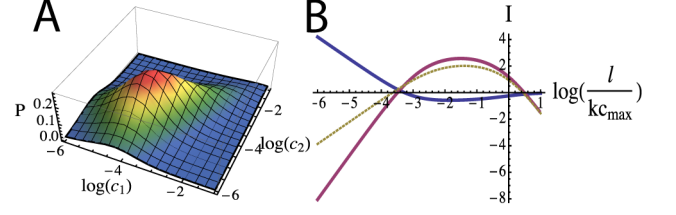


FIG. 3. Optimal input distribution and channel capacity. (A) The optimal input distribution for $l = 10^{-4}$ and $c_{min} = 10^{-3}$ as a function of log TF concentration. (B) The channel capacity in bits for the interacting and non-interacting case of two TFs (dashed-brown and solid-red curves respectively) as a function of logarithm of rescaled l . The y-offset is arbitrary. Blue curve denotes their difference. At biologically relevant $l = 10^{-4}$, interacting TFs have higher channel capacity.

arable, $P_{TF}^*(c_1, c_2) \neq P_1^*(c_1)P_2^*(c_2)$. With an entangled distribution the system can explore degrees of freedom not present with two independent input distributions. In Fig. 4, we plot the log likelihood of observing joint concentration (c_1, c_2) compared to observing c_1 and c_2 independently from their marginal distributions. $R = \log_{10} \frac{P_{TF}^*(c_1, c_2)}{P_1^*(c_1)P_2^*(c_2)}$, where $P_1^*(c_1) = \int dc_2 P_{TF}^*(c_1, c_2)$ is the marginal distribution of TF₁, with a similar expression for TF₂.

Fig. 4 implies that the two TFs are no longer passive and in fact interact with each other. It is ~ 10 times less likely to observe one TF at a high concentration and the other at a low concentration simultaneously, compared to what is expected if they were independent. Similarly, it is ~ 10 times more likely to observe high concentrations of one TF if the concentration of the other is also high. This suggests that one TF positively regulates the other (see motif in Fig. 4). We assume that the noise introduced in TF concentration from cross-regulation is small compared to the noise in the read-out during expression. This is reasonable if the physical mechanism of cross-talk is different, or the cell is allowed longer integration times for TF regulation than that of the target gene.

V. BIOLOGICAL EXAMPLES

Where does a biological system lie in the abstract parameter space sketched above? As noted, we have rescaled time so that $k = 1$, and measured concentration in units of $c_{max} = 1$. With the assumption of fast integration time, the only parameters left are the off-rate l and c_{min} . In a real cell, we expect a absolute maximum of roughly 1000 TF molecules (or a dynamic range of 1-1000 TF molecules) in a volume of $\sim 1 \mu m^3$ [23]. Hence, the minimum allowed concentration is $c_{min} = 10^{-3}$. A typical equilibrium constant of TF binding to DNA is $K_{eq} \sim 10^{10}$ [22]. Putting all this together, we find $l \sim 10^{-4}$. It is possible then that a real biological regula-

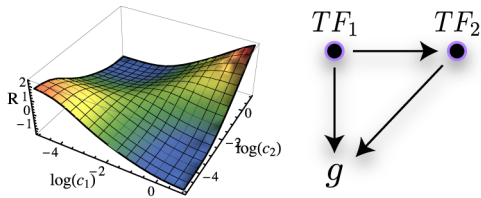


FIG. 4. TF upstream cross-regulation. The log likelihood of observing TF concentration (c_1, c_2) compared to what is expected from independent distributions. It is much more likely to observe both TFs at either low or high concentrations together. This suggests that one TF positively regulates the other, feed-forward motif (right).

tory system can transmit more information by incorporating overlapping binding sites and an upstream positive feedback between the TFs.

Concentration-dependent transcription regulation is particularly important at the developmental stage. Concentration of morphogens dictate cell fate, for example, resulting in patterning of the *Drosophila* embryo along the dorsoventral axis [25]. It is likely that the embryo has optimized information transmission to ensure accurate patterning and later development. Gene regulation using a combination of transcription factors is also a common theme in development (see for example [26]).

Xu et al. [24] have observed the feed-forward motif in regulation of gene *Race* in the *Drosophila* embryo. They report that intracellular protein Smads sets the expression level of *zerknüllt* (*zen*), then Smads in combination with *zen* (two-fold input) directly activate *Race*. Analysis of binding site of Smads and *zen* reveals slight overlaps, and experiments indicate that one protein facilitates binding of the other to the enhancer. This interaction can result in a similar positive correlation coefficient in TF concentration estimates derived in the toy model above. The feed-forward motif then is predicted by the framework outlined in last section. The previously proposed suggestion [24] that feed-forward motif increases sensitivity to the input signal does not explain why the target is regulated by both the initial input and the target transcription factor. Proposed dynamical features associated with the feed-forward loop [9, 27] do not explain the need for overlapping binding sites and TF interactions at binding.

Another example of a feed-forward motif coupled to binding interactions is the joint-regulation of *even-skipped* (*eve*) stripe 2 by *bicoid* (*bcd*) and *hunchback* (*hb*). Small et al. [3] report cooperative binding interactions between *bcd* and *hb* and a clustering of their binding sites in the promoter region. Upstream, *bcd* positively regulates transcription of *hb*. Similarly, Ip et al. [4] have observed joint-activation of gene *snail* (*sna*) by *twist* (*twi*) and *dorsal* (*dl*), which also exhibit cooperative binding

interactions. *dl* directly regulates transcription of two upstream.

More generally, other forms of cross-talk besides transcriptional regulation can be used. For instance, in regulation of anaerobic respiration in *E. coli*, regulators NarP and NarL are jointly-regulated through phosphorylation by histidine kinase NarX. NarL is also phosphorylated by kinase NarX. Downstream, NarP and NarL share the same DNA binding site [28]. It is not, however, clear if optimizing channel capacity is relevant for this system. It is also possible that interference is implemented using other schemes than DNA binding, for example, through cooperative interactions of signaling molecules with scaffold proteins [29].

VI. DISCUSSION

Using a simple model, we have shown that TF interaction at overlapping binding sites plus upstream cross-regulation can enhance information transmission compared to non-interacting TFs. We motivated the frequent observation of the feed-forward motif ending in overlapping binding sites in developmental gene networks. Although the feed-forward motif has been proposed before for optimizing information transmission in regulatory networks [16], we emphasize that our approach is fundamentally different, since it stems from correlated binding noise, and physically requires existence of TF interactions at the binding level. This is indeed what is experimentally observed in the three examples discussed above. We also hope to convey the importance of careful analysis of correlated noise in biological systems. The conventional approach of treating intracellular noise as uncorrelated fluctuations might not reveal intricate connections in the complex networks involved in regulation, transcription, and translation.

A myriad of logical regulatory circuits have been proposed through use of overlapping binding sites and interacting TFs [23, 30]. It is worthwhile to see if the upstream TF regulatory network of these systems can be correctly predicted from the overlap/interactions using the methodology outlined above. This will shed light on whether optimization of channel capacity is biologically relevant for gene regulation. Such analysis requires knowledge of the binding noise, which can be obtained in two ways. One is a bioinformatics approach, where the binding sequence of each TF is examined for overlap. A physical model is then required to make predictions on the nature of the noise from the overlap (a toy example of which was presented here) and possibly protein-protein interactions. The other approach requires measurement of the noise directly using single-molecule techniques [31], and use in Eq.(6) to predict upstream TF cross-regulation.

ACKNOWLEDGMENTS

The author thanks Boris Shraiman for helpful discussions and critical reading of the manuscript, and Bill Bialek for introduction to the subject. This research was supported in part by the National Science Foundation under Grant No. NSF PHY05-51164.

-
- [1] Wolpert L, et al. (2006) Principles of Development (Oxford Univ Press), 3rd Ed.
 - [2] Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* **25**:1-47.
 - [3] Small S, Blair A, Levine M (1992) Regulation of even-skipped stripe 2 in the Drosophila embryo. *The EMBO Journal* **11**:4047-4057.
 - [4] Ip YT, Park RE, Kosman D, Yazdanbakhsh K, Levine M (1992) dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo. *Genes Dev* **6**:1518-1530.
 - [5] Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* **297**:1183-1186.
 - [6] Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* **99**:12795-12800.
 - [7] Paulsson J (2004) Summing up the noise in gene networks. *Nature* **427**:415-418.
 - [8] Tkacik G, Callan CG, Bialek W (2008) Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci USA* **105**:12265-12270.
 - [9] Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31**:64-68.
 - [10] Lee TI, et al. (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science* **298**:799-804.
 - [11] Tkacik G, Callan CG, Bialek W (2008) Information capacity of genetic regulatory elements. *Phys Rev E* **78**:11910.
 - [12] Gregor T, Tank DW, Wieschaus EF, Bialek W (2007) Probing the limits to positional information. *Cell* **130**:153-164.
 - [13] Raser JM, O'shea EK (2004) Control of stochasticity in eukaryotic gene expression. *Science* **304**:1811-1814.
 - [14] Newman JR, et al. (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**:840-846.
 - [15] Rosenfeld N, et al. (2005) Gene regulation at the single-cell level. *Science* **307**:1962-1965.
 - [16] Walczak AM, Tkacik G, Bialek W (2010) Optimizing information flow in small genetic networks. II. Feed-forward interactions. *Phys Rev E* **81**:041905.
 - [17] Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci USA* **102**:10040-10045.
 - [18] Kubo R (1966) The fluctuation-dissipation theorem. *Rep Prog Phys* **29**:255-284.
 - [19] Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* **20**:193-219.
 - [20] Shannon CE (1949) Communication in the presence of noise. *Proc IRE* **37**:10-21.
 - [21] Cover TM, Thomas JA (1991) Elements of Information Theory (Wiley, New York).
 - [22] Bintu L, et al. (2005) Transcriptional regulation by the numbers: Applications. *Curr Opin Genet Dev* **15**:125-135.
 - [23] Buchler NE, Gerland U, Hwa T (2003) On schemes of combinatorial transcription logic. *Proc Natl Acad Sci USA* **100**:5136-5141.
 - [24] Xu M, Kirov N, Rushlow C (2005) Peak levels of BMP in the Drosophila embryo control target genes by a feed-forward mechanism. *Development* **132**:1637-1647.
 - [25] Morisato D, Anderson KV (1995) Signaling pathways that establish the dorsalventral pattern of the Drosophila embryo. *Annu. Rev. Genet.* **29**:371-99.
 - [26] Howard ML, Davidson EH (2004). cis-Regulatory control circuits in development. *Dev. Biol.* **271**:109-118.
 - [27] Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci USA* **100**:11980-11985.
 - [28] Stewart V (2003) Biochemical Society Special Lecture. Nitrate- and nitrite-responsive sensors NarX and NarQ of proteobacteria. *Biochem. Soc. Trans.* **31**:1-10.
 - [29] Good MC, Zalatan JG, Lim WA (2011) Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science* **332**:680-686.
 - [30] Hermesen R, Tans S, tenWolde P-R (2006) Transcriptional Regulation by Competing Transcription Factor Modules. *PLoS Comp. Biol.* **2**:e164. s
 - [31] Yufang Wang, Ling Guo, Ido Golding, Edward C. Cox, and N. P. Ong Quantitative Transcription Factor Binding Kinetics at the Single-Molecule Level. *Biophysical Journal* Volume **96**:609-620.